

## Transcriptome characterization of gonadal sex differentiation in Pacific bluefin tuna, *Thunnus orientalis* (Temminck et Schlegel)

メタデータ	言語: English 出版者: 公開日: 2024-01-12 キーワード (Ja): キーワード (En): 作成者: 林田, 貴雄, 相馬, 智史, 中村, 洋路, 樋口, 健太郎, 風藤, 行紀, 玄, 浩一郎 メールアドレス: 所属: 水産研究・教育機構, 水産研究・教育機構, 水産研究・教育機構, 水産研究・教育機構, 水産研究・教育機構, 水産研究・教育機構
URL	<a href="https://fra.repo.nii.ac.jp/records/2000067">https://fra.repo.nii.ac.jp/records/2000067</a>



## OPEN Transcriptome characterization of gonadal sex differentiation in Pacific bluefin tuna, *Thunnus orientalis* (Temminck et Schlegel)

Takao Hayashida<sup>1,2✉</sup>, Satoshi Soma<sup>3</sup>, Yoji Nakamura<sup>3</sup>, Kentaro Higuchi<sup>1,4</sup>, Yukinori Kazeto<sup>4</sup> & Koichiro Gen<sup>1</sup>

Tunas (genus *Thunnus*) are one of the most ecologically and commercially important fish worldwide. To establish a biological basis for reproduction in this globally essential species, we have recently studied crucial reproductive aspects of the Pacific bluefin tuna (*T. orientalis*; PBT), as a model of tuna species, based on our closed-cycle aquaculture technology. In this study, we clarified the global expression profile of the genes regulating gonadal sex differentiation in PBT, as this developmental process is vital to sexual reproduction. Based on the results of our comparative (RNA-sequencing) and temporal (qRT-PCR) transcriptome analyses using the updated genome dataset, we propose the molecular mechanisms of gonadal sex differentiation in PBT. In female gonads, *foxl2* and *cyp19a1a* (coding aromatase) are expressed at the onset of sex differentiation. Active aromatase-mediated estrogen biosynthesis, which includes positive regulation of *cyp19a1a* expression by *Foxl2*, induces ovarian differentiation. By contrast, *dmrt1* and *gsdf* are upregulated in differentiating male gonads lacking active estrogen synthesis. *Dmrt1* and *Gsdf* would mainly promote testicular differentiation. Furthermore, androgen biosynthesis is upregulated in differentiating male gonad. Endogenous androgens may also be vital to testicular differentiation. This study provides the first comprehensive data clarifying the molecular basis for gonadal sex differentiation in tunas.

Tunas (*Thunnus* spp.) are widely distributed in the global ocean and are top predators in marine ecosystems. Tuna, specifically bluefin tuna, has a high market demand and value and is of substantial commercial interest to fisheries and aquaculture operations worldwide. Despite the ecological and commercial importance of tuna, the mechanisms by which its reproductive system develops remain poorly understood, mainly owing to the difficulty in obtaining artificial seedlings that can be used for experiments and their rearing<sup>1</sup>. A detailed understanding of the biological basis for tuna reproduction is necessary to achieve several goals in fisheries and aquaculture operations, such as effective assessments of population fluctuations in wild-stock management<sup>2</sup> and efficient broodstock management in captivity to obtain high-quality fertilized eggs<sup>3,4</sup>. Hence, we recently studied the vital reproductive aspects of Pacific bluefin tuna (*T. orientalis*; hereafter, PBT), as a model of tuna species, based on our closed-cycle aquaculture technology<sup>1,5–8</sup>.

Gonadal sex differentiation is crucial in vertebrate sexual reproduction. In gonochoristic fishes (many species of fish, including PBT<sup>6</sup>, are gonochoristic<sup>9</sup>), gonadal primordia are formed at the early embryonic stage and subsequently differentiate into ovaries or testes under the control of various genes and factors<sup>10</sup>. Estrogens are critical factors in fish ovarian differentiation<sup>10–12</sup>. *Cyp19a1a* (cytochrome P450 aromatase, P450-arom; hereafter, aromatase) is expressed mainly in the female gonads at the onset of sex differentiation, and estrogens synthesized by aromatase directly induce ovarian differentiation<sup>10–12</sup>. The transcription factor forkhead box protein L2 (*Foxl2*) is also a major player in active estrogen biosynthesis as it facilitates *aromatase* expression<sup>11,12</sup>. By contrast, the physiological roles of androgens in testicular differentiation vary among fish species. For instance, androgen

<sup>1</sup>Nagasaki Field Station, Fisheries Technology Institute, Japan Fisheries Research and Education Agency, 1551-8 Taira-machi, Nagasaki, Nagasaki 851-2213, Japan. <sup>2</sup>Graduate School of Marine Science and Technology, Tokyo University of Marine Science and Technology, 4-5-7 Konan, Minato-ku, Tokyo 108-8477, Japan. <sup>3</sup>Yokohama Field Station, Fisheries Resources Institute, Japan Fisheries Research and Education Agency, 2-12-4 Fuku-ura, Yokohama, Kanagawa 236-8648, Japan. <sup>4</sup>Minamiizu Field Station, Fisheries Technology Institute, Japan Fisheries Research and Education Agency, 183-2 Minamiizu, Kamo, Shizuoka 415-0156, Japan. ✉email: hayashida\_takao43@fra.go.jp

production is absent in the male gonads of Nile tilapia (*Oreochromis niloticus*) during testicular differentiation<sup>13,14</sup>. Moreover, androgens are actively synthesized in the differentiating male gonads of rainbow trout (*Oncorhynchus mykiss*)<sup>15,16</sup>. Doublesex and mab-3 related transcription factor 1 (Dmrt1) is highly conserved and orchestrates testicular differentiation in various different vertebrates<sup>10</sup>. Dmrt1 facilitates the expression of genes regulating testicular differentiation, such as *amh* (anti-Müllerian hormone) in zebrafish (*Danio rerio*)<sup>17</sup> and *sox9b* (SRY-box transcription factor 9b) in Nile tilapia<sup>18</sup>. Dmrt1 may also directly block estrogen production as reported in Nile tilapia<sup>19</sup>. The gonadal soma-derived factor (Gsd) also induces testicular differentiation in fish<sup>10</sup>. *Gsd* upregulation in differentiating male gonads has been confirmed for various fish species<sup>10</sup>. The vital role of *Gsd* in testicular differentiation has been demonstrated through loss- and/or gain-of-function studies in several fish species including Nile tilapia<sup>20</sup> and medaka (*Oryzias latipes*)<sup>21</sup>.

We recently demonstrated the timing and morphological process of gonadal sex differentiation in PBT<sup>6</sup>. Furthermore, no difference in germ cell proliferation and differentiation between sexes during gonadal sex differentiation was observed<sup>6</sup>. We also showed the vital role of aromatase-mediated estrogen biosynthesis in ovarian differentiation in PBT. Aromatase inhibitor (AI) administration resulted in sex reversal of genotypic female to phenotypic male<sup>1</sup>. We confirmed testicular differentiation of the AI-treated genotypic female PBT gonads at the molecular level through histological observation, gene expression analysis, and serum sex steroid measurement<sup>1</sup>. However, the detailed molecular mechanisms of gonadal sex differentiation in general, and testicular differentiation in particular, remain unclear. A detailed understanding of the gonadal sex differentiation mechanisms will provide a fundamental basis for developing reproductive biotechnology in aquaculture, specifically sex-manipulation technologies<sup>22</sup>.

Application of next-generation sequencing (NGS) technology has recently become increasingly common in the field of aquaculture and fisheries. High-throughput RNA sequencing (RNA-Seq) technology is effective in studying the molecular mechanisms underlying the various biological characteristics of non-model and model species and has high reproducibility. The objective of the present study was to elucidate the molecular mechanisms of gonadal sex differentiation in PBT using NGS technologies. Our research group published the first-draft PBT genome in 2013<sup>23</sup> and improved its assembly in 2019<sup>25</sup> and again in 2021<sup>25</sup>. In this study, we initially polished the 2021 version of our PBT draft genome assembly<sup>25</sup> and predicted the protein coding sequences using the refined genome. Using the predicted gene dataset as a reference, we then examined the global gene expression profiles of gonadal sex differentiation in PBT via comparative transcriptome analysis based on RNA-Seq technology. Furthermore, we examined the temporal expression profiles of the candidate genes controlling gonadal sex differentiation using quantitative real-time reverse transcription PCR (qRT-PCR).

## Results

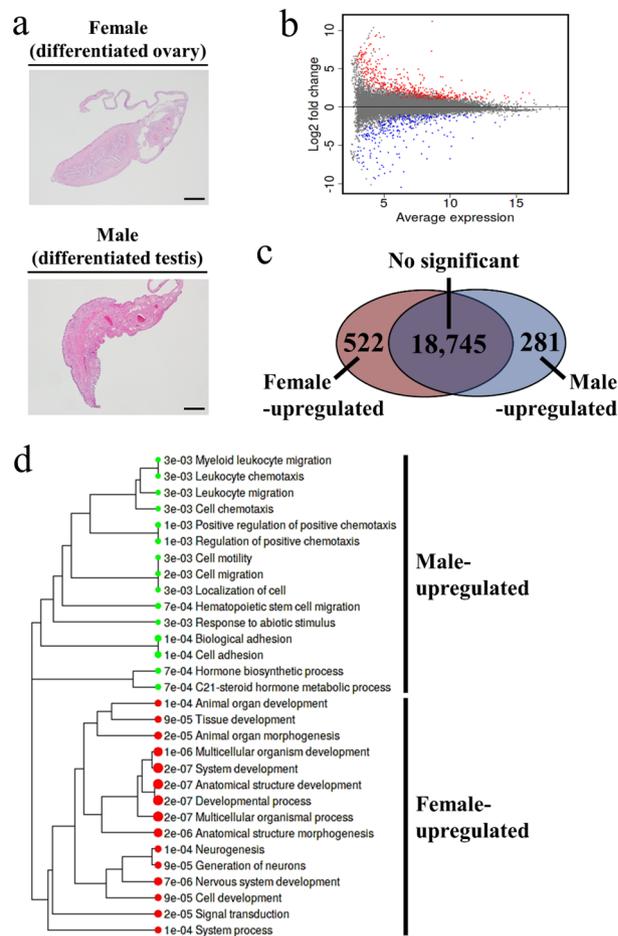
**Reference gene dataset construction using a polished draft genome.** The 2021 version of the PBT draft genome<sup>25</sup> was rearranged from 948 to 580 scaffolds by polishing the sequence and eliminating redundancy. The completeness score obtained via Benchmarking Universal Single-Copy Orthologs (BUSCO) assessment, i.e., the reference ortholog capture ratio of the older version was 98.2%, of which 95.7% were singletons and 2.5% were duplicated<sup>25</sup>. This ratio was 98.4% for the current version, of which 97.7% were singletons and 0.7% were duplicated. Few of the doubly predicted orthologs in the previous version (~2% of 3640 reference genes)<sup>25</sup> were predicted as singletons by removing redundant scaffolds. We predicted 30,156 protein-coding sequences from the current PBT genome (Supplementary Data 1). Of these, ~90% resembled those corresponding to the zebrafish and medaka protein sequences (see Supplementary Table S1 online).

**RNA-sequencing and comparative transcriptome analyses.** Table 1 summarizes the RNA sequencing results. Sequencing of the cDNA libraries yielded 75–102 million paired-end reads (75 bp). Over 99% of the sequenced reads were clean following quality control. Clean reads (89.9–90.7%) were read-mapped to the reference sequences.

The PBT gonads at the sex-differentiated stage (Fig. 1a) were subjected to comparative transcriptome analysis (Fig. 1b,c). A total of 522 and 281 genes were significantly upregulated in the females and males, respectively ( $P < 0.05$ , see Supplementary Table S2 online). Both sexes shared the expression of 18,745 genes ( $P > 0.05$ , see Supplementary Table S3 online). Figure 1d and Supplementary Tables S4–S5 show the top 15 Gene Ontology (GO) biological process terms enriched in each sex from the genes differentially expressed between sexes. In the females, several cell/tissue development-related terms, including ‘animal organ development’ (GO: 0,048,513), ‘tissue development’ (GO: 0,009,888), ‘animal organ morphogenesis’ (GO: 0,009,887), ‘multicellular organism

Sample	RNA integrity no. (RIN)	No. reads	No. clean reads	No. clean bases (Gb)	GC content (%)	Total No. mapped reads (%)
Female-1	9.2	79,685,638	79,682,070	5.21	47.6	90.7
Female-2	8.9	75,238,374	75,234,984	4.92	47.3	90.7
Female-3	8.3	102,356,556	102,352,454	6.70	47.0	90.3
Male-1	7.3	81,855,526	81,852,002	5.36	46.9	89.9
Male-2	8.2	77,421,316	77,417,522	5.07	46.7	90.6
Male-3	7.7	78,289,102	78,285,284	5.12	46.9	90.1

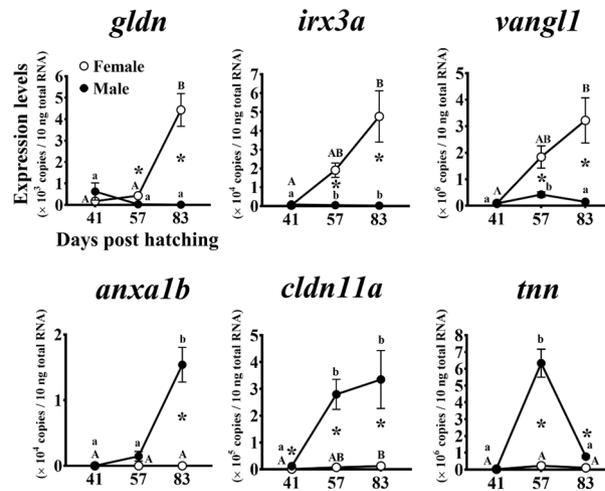
**Table 1.** Results of RNA sequencing.



**Figure 1.** Summary of comparative transcriptome analysis. (a) Histological micrographs of sex-differentiated stage Pacific bluefin tuna gonads subjected to RNA sequencing. Bars = 200  $\mu$ m. (b) MA plot showing differentially expressed genes. Significantly upregulated ( $P < 0.05$ ) genes in females and males are indicated by red and blue dots, respectively. Genes not significantly different between sexes ( $P > 0.05$ ) are indicated by gray dots. (c) Venn diagram showing numbers of commonly ( $P > 0.05$ ) and differentially ( $P < 0.05$ ) expressed genes between sexes. (d) Hierarchical clustering tree showing the top 15 enriched Gene Ontology (GO) biological process terms in upregulated genes for each sex. The circle size and number on left side of each GO terms indicate false discovery rate adjusted  $P$ -values.

development' (GO: 0,007,275), 'system development' (GO: 0,048,731), 'anatomical structure development' (GO: 0,048,856), 'developmental process' (GO: 0,032,502), 'multicellular organismal process' (GO: 0,032,501), 'anatomical structure morphogenesis' (GO: 0,009,653), 'cell development' (GO: 0,048,468), and 'system development' (GO: 0,003,008) were detected. In the males, structural development-related terms such as 'cell motility' (GO: 0,048,870), 'cell migration' (GO: 0,016,477), 'localization of cell' (GO: 0,051,674), 'biological adhesion' (GO: 0,022,610), and 'cell adhesion' (GO: 0,007,155) were detected. The steroidogenesis-related terms 'hormone biosynthetic process' (GO: 0,042,446) and 'C21-steroid hormone metabolic process' (GO: 0,008,207) were also detected in the males.

The expression profile of the genes enriched in several GO terms was exemplarily validated using qRT-PCR (Fig. 2). The qRT-PCR targeted three upregulated genes for each sex: *gldn* (*gliomedin*; contig No. g2699), *irx3a* (*iroquois homeobox 3a*; contig No. g3224), and *vangl1* (*VANGL planar cell polarity protein 1*; contig No. g5686) in female, and *anxa1b* (*annexin A1b*; contig No. g5180), *cldn11a* (*claudin 11a*; contig No. g10146), and *tnn* (*tenascin N*; contig No. g11982) in males (see Supplementary Table S4–S5 online). Furthermore, their temporal expression during gonadal sex differentiation was identified. In the females, the *gldn*, *irx3a*, and *vangl1* expressions significantly increased from the morphologically sex-undifferentiated (41 days post hatching) to the differentiated (83 dph) stages ( $P < 0.05$ ). However, in the males, the *gldn*, *irx3a*, and *vangl1* expressions remained low until the differentiated stage. The *gldn*, *irx3a*, and *vangl1* expressions were significantly higher in the females than in the males at the differentiating (57 dph) and differentiated (83 dph) stages ( $P < 0.05$ ). In the males, the *anxa1b* and *cldn11a* expressions significantly increased from the undifferentiated (41 dph) to the differentiated (83 dph) stages ( $P < 0.05$ ). The *cldn11a* and *tnn* expressions significantly increased from the undifferentiated (41 dph) to the differentiating (57 dph) stages ( $P < 0.05$ ). The *anxa1b*, *cldn11a*, and *tnn* expressions remained low until the differentiated stage in the female. The *anxa1b*, *cldn11a*, and *tnn* expressions were significantly higher in the



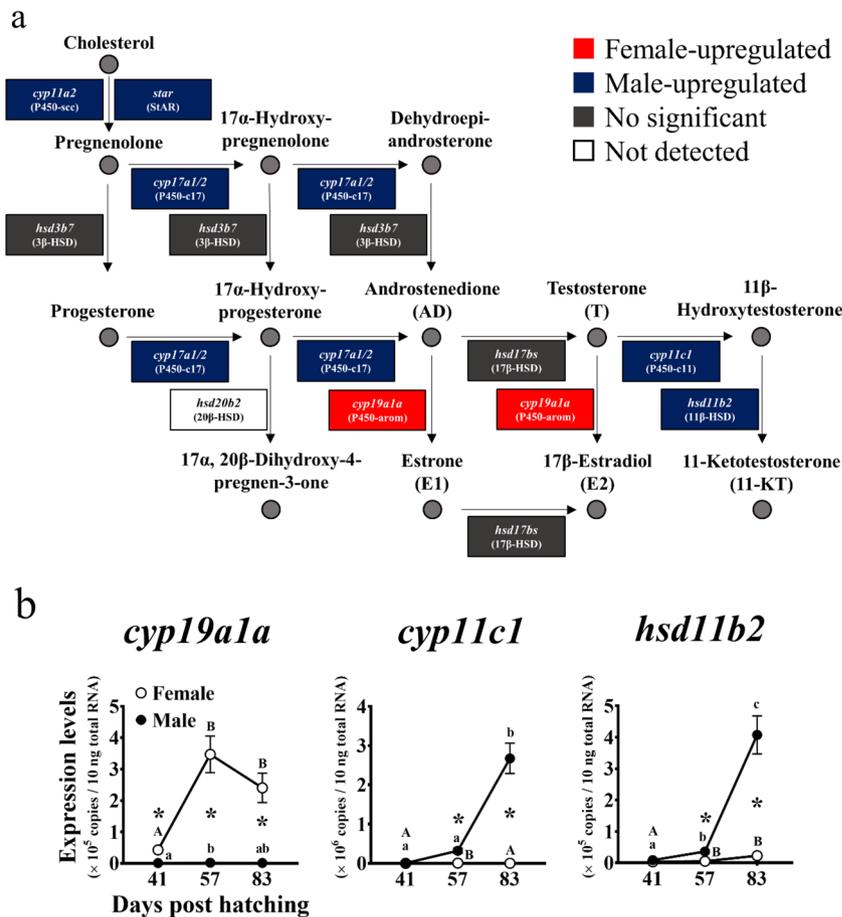
**Figure 2.** Temporal expression patterns of genes encoding unreported factors in fish sex differentiation during gonadal sex differentiation in Pacific bluefin tuna. Quantitative real-time reverse transcription (qRT-PCR) targeted three genes for each sex which were enriched in multiple Gene Ontology (GO) pathways (*gldn*, *irx3a*, and *vangl1* in females, and *anxa1b*, *cldn11a*, and *ttn* in males; see Supplementary Table S4–S5 online). Total RNA extracted from the gonads of each sex at morphologically sex-undifferentiated (41 days post-hatching (dph)), differentiating (57 dph), and differentiated (83 dph) stages<sup>6</sup> were subjected to qRT-PCR. Open and closed circles indicate genotypic females and males, respectively. Bars indicate the mean  $\pm$  standard error of the mean (SEM) ( $n=3$  fish). Different letters indicate significant differences ( $P<0.05$ , one-way ANOVA followed by Tukey's multiple comparisons test). Asterisks indicate significant differences between sexes at each age ( $P<0.05$ , Welch's  $t$ -test).

males than in the females at the differentiated stage (83 dph;  $P<0.05$ ). The *cldn11a* and *ttn* expressions were also significantly higher in the males than in the females at the differentiating stage (57 dph;  $P<0.05$ ). The *cldn11a* expression was significantly higher in the males than in the females at the undifferentiated stage (41 dph;  $P<0.05$ ).

**Sex steroid synthesis-related gene expression patterns.** A comparative transcriptome analysis disclosed sexual dimorphism in sex steroid synthesis-related gene expression (Fig. 3a). *Star* (steroidogenic acute regulatory protein, StAR; contig No. g8869), *cyp11a2* (cytochrome P450 cholesterol side-chain cleavage enzyme, P450-scc; contig No. g1984), *cyp17a1* (cytochrome P450 17 $\alpha$ -hydroxylase/C17-C20 lyase, P450-c17; contig No. g4434), *cyp17a2* (P450-c17; contig No. g26576), *cyp11c1* (cytochrome P450 family 11 subfamily C polypeptide 1, P450-c11; contig No. g16537), and *hsd11b2* (11 $\beta$ -hydroxysteroid dehydrogenase, 11 $\beta$ -HSD; contig No. g28927) were significantly upregulated in the males ( $P<0.05$ ). *Aromatase* (contig No. g2700) was significantly upregulated in the females ( $P<0.05$ ). There were no significant differences between sexes in terms of their *hsd3b7* (3 $\beta$ -hydroxysteroid dehydrogenase, 3 $\beta$ -HSD; contig No. g17096) and *hsd17bs* (17 $\beta$ -hydroxysteroid dehydrogenase, 17 $\beta$ -HSDs; see Supplementary Fig. S1 online) expression levels ( $P>0.05$ ). *Hsd20b2* (20 $\beta$ -hydroxysteroid dehydrogenase, 20 $\beta$ -HSD; contig No. g14056) was not expressed in either sex. Furthermore, there were no significant differences in the expressions of estrogen receptor encoding genes, *esr1* (*estrogen receptor 1*; contig No. g5109), *esr2a* (*estrogen receptor 2a*; contig No. g8921), and *esr2b* (*estrogen receptor 2b*; contig No. g22496) between sexes ( $P>0.05$ ; see Supplementary Table S3 and Fig. S2 online).

The qRT-PCR identified the temporal expression of the genes encoding estrogen and androgen biosynthesis-related enzymes (*aromatase*, *cyp11c1*, and *hsd11b2*) during gonadal sex differentiation (Fig. 3b). *Aromatase* expression data have been previously reported<sup>1</sup>. In the females, the *aromatase* expression significantly increased from the morphologically sex-undifferentiated (41 days post-hatching (dph)) to the differentiating (57 dph) stages ( $P<0.05$ ) and remained high until the differentiated stage (83 dph). In the males, however, the *aromatase* expression remained low until the differentiated stage. The *Aromatase* expression was significantly higher in the females than in the males at all stages (41–83 dph;  $P<0.05$ ). In the males, the *cyp11c1* expression significantly increased from the differentiating (57 dph) to the differentiated (83 dph) stages ( $P<0.05$ ). In the females, however, it remained low until the differentiated stage. In the males, the *hsd11b2* expression significantly increased from the undifferentiated (41 dph) to the differentiating (57 dph) stages ( $P<0.05$ ) and increased further until the differentiated (83 dph) stage ( $P<0.05$ ). In the females, however, it remained low until the differentiated stage. The *cyp11c1* and *hsd11b2* expression levels were significantly higher in the males than in the females at the differentiating (57 dph) and differentiated (83 dph) stages ( $P<0.05$ ).

**Sex differentiation-related gene expression patterns.** A comparative transcriptome analysis revealed sexual dimorphism in sex differentiation-related gene expression (Fig. 4a). *Foxl2* (contig No. g17853) and *aromatase* were significantly upregulated in the females ( $P<0.05$ ), whereas *dmrt1* (contig No. g7920) and *gsdf* (contig No. g4772) were significantly upregulated in the males ( $P<0.05$ ). There were no significant

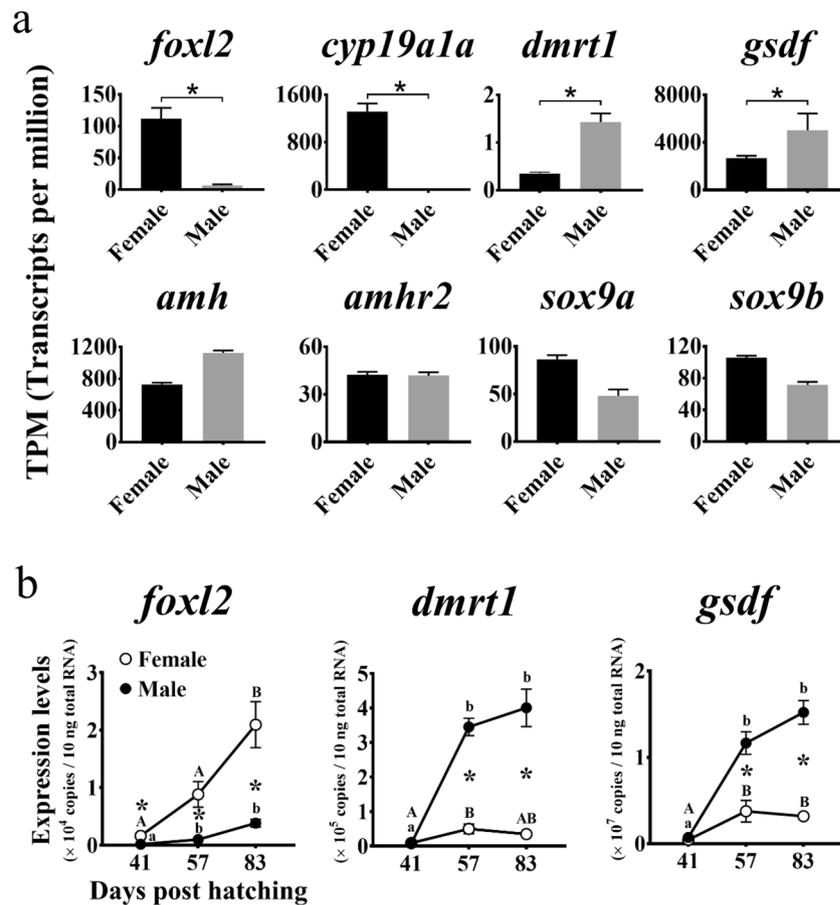


**Figure 3.** (a) Expression patterns of sex steroid biosynthesis-related genes in gonads of Pacific bluefin tuna at the sex differentiated stage. The diagram shows gonadal sex steroid biosynthesis pathways in fish<sup>26,27</sup>. Upregulated genes in females and males are indicated by red and blue boxes, respectively ( $P < 0.05$ ). Genes not significantly different between sexes are indicated by gray boxes ( $P > 0.05$ ). The gene with no detectable expression is indicated by a white box. Expression patterns of genes encoding 17 $\beta$ -hydroxysteroid dehydrogenases (17 $\beta$ -HSDs) are shown in Supplementary Fig. S1. (b) Temporal expression patterns of genes encoding key enzymes in estrogen and androgen biosynthesis during gonadal sex differentiation in Pacific bluefin tuna. Total RNA extracted from gonads in each sex at morphologically sex-undifferentiated (41 days post-hatching (dph)), differentiating (57 dph), and differentiated (83 dph) stages<sup>6</sup> subjected to quantitative real-time reverse transcription PCR. Open and closed circles indicate genotypic females and males, respectively. Bars indicate mean  $\pm$  standard error of the mean (SEM) ( $n = 3$  fish). Different letters indicate significant differences ( $P < 0.05$ , one-way ANOVA followed by Tukey's multiple comparisons test). Asterisks indicate significant differences between sexes at each age ( $P < 0.05$ , Welch's  $t$ -test). *Cyp19a1a* expression data were previously reported<sup>1</sup>.

differences between sexes in terms of their *amh* (contig No. g11182), *amhr2* (anti-Müllerian hormone receptor type 2; contig No. g370), *sox9a* (SRY-box transcription factor 9a; contig No. g15913), and *sox9b* (contig No. g14928) expression levels ( $P > 0.05$ ).

The qRT-PCR identified the temporal expression of the sex differentiation-related genes with sexually dimorphic expression during gonadal sex differentiation (Fig. 4b). In the females, the *foxl2* expression significantly increased from the differentiating (57 dph) to the differentiated (83 dph) stages ( $P < 0.05$ ). In the males, however, the *foxl2* expression remained low until the differentiated stage. The *Foxl2* expression was significantly higher in the females than in the males at all stages (41–83 dph;  $P < 0.05$ ). In the males, the *dmrt1* and *gsdf* expression levels significantly increased from the undifferentiated (41 dph) to the differentiating (57 dph) stages ( $P < 0.05$ ) and remained high until the differentiated stage (83 dph). In the females, however, the *dmrt1* and *gsdf* expression levels remained low until the differentiated stage. The *dmrt1* and *gsdf* expression levels were significantly higher in the males than in the females at the differentiating (57 dph) and the differentiated (83 dph) stages ( $P < 0.05$ ).

**Sex steroid synthesis- and sex differentiation-related gene expression in sex-reversed gonads.** Figure 5 shows the sex steroid synthesis- and sex differentiation-related gene expression levels in sex-reversed gonads (genotypic female/phenotypic male) artificially induced by the AI treatment<sup>1</sup>. The

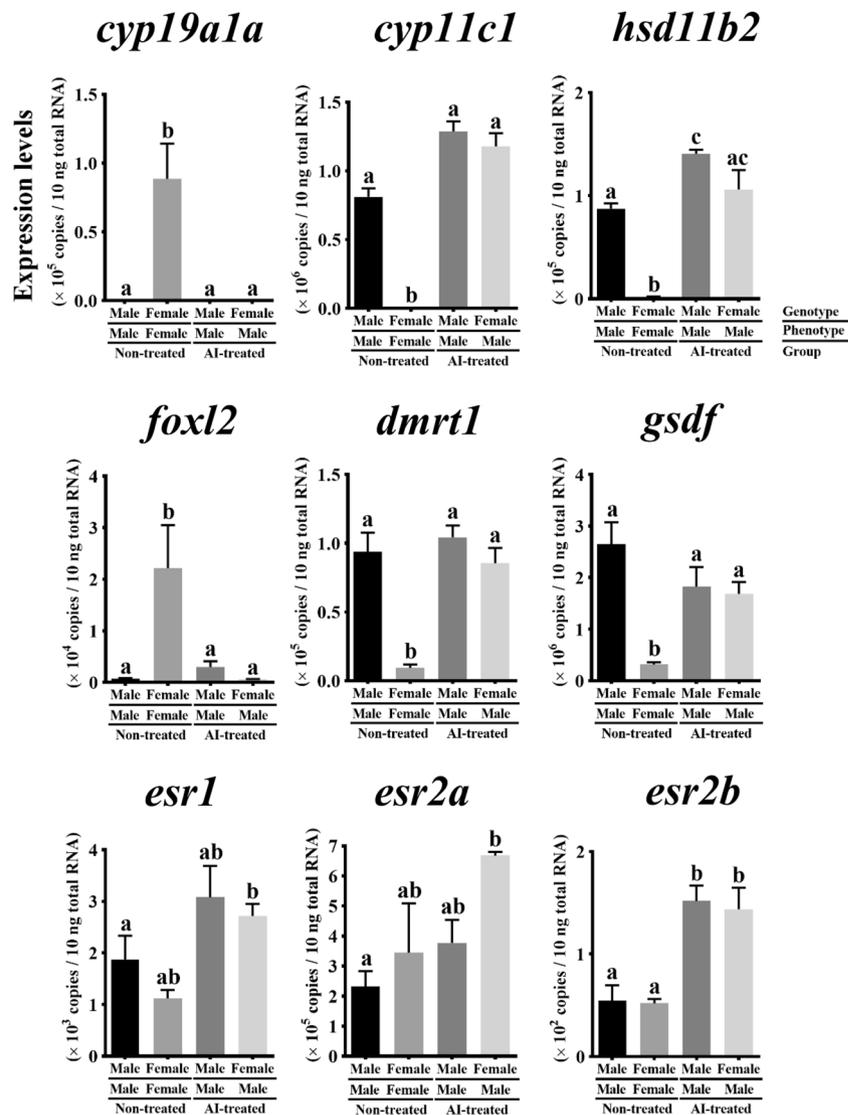


**Figure 4.** (a) Sex differentiation-related gene expression patterns in Pacific bluefin tuna gonads at the sex-differentiated stage. Expression levels correspond to transcripts per million (TPM). Statistical differences indicated by asterisks are false discovery rate-adjusted  $P$ -values ( $P < 0.05$ ). (b) Temporal expression patterns of genes identified via comparative transcriptome analysis as candidate gonadal sex differentiation regulators in Pacific bluefin tuna. Total RNA extracted from the gonads of each sex at morphologically sex-undifferentiated (41 days post-hatching (dph)), differentiating (57 dph), and differentiated (83 dph) stages<sup>6</sup> were subjected to quantitative real-time reverse transcription PCR. Open and closed circles indicate genotypic females and males, respectively. Bars indicate mean  $\pm$  standard error of the mean (SEM) ( $n = 3$  fish). Different letters indicate significant differences ( $P < 0.05$ , one-way ANOVA followed by Tukey's multiple comparisons test). Asterisks indicate significant differences between sexes at each age ( $P < 0.05$ , Welch's  $t$ -test).

*aromatase* and *foxl2* expression levels were significantly higher in the genotypic/phenotypic females than the others ( $P < 0.05$ ). The *cyp11c1*, *hsd11b2*, *dmrt1*, and *gsdf* expression levels in the genotypic female/phenotypic males were comparable to those in the genotypic/phenotypic males in both the AI-treated and untreated groups ( $P > 0.05$ ). Nevertheless, the *cyp11c1*, *hsd11b2*, *dmrt1*, and *gsdf* expression levels were significantly lower in the genotypic/phenotypic females than the others ( $P < 0.05$ ). No significant differences in the *esr1*, *esr2a*, and *esr2b* expressions were observed between the sexes in each treated group ( $P > 0.05$ ). The expression levels of the estrogen receptors were higher in the AI-treated group than in the non-treated group. Furthermore, *esr2b* expression were significantly higher in the AI-treated group than those in the non-treated group ( $P < 0.05$ ).

## Discussion

In this study, we initially updated our draft genome assembly for PBT. The completeness scores indicate that the current polished genome was slightly improved compared to the previous version<sup>25</sup>. Thereafter, we transcriptomically characterized gonadal sex differentiation in PBT using a reference gene dataset predicted from the polished draft genome. A comparative transcriptome analysis based on RNA-Seq identified 19,548 genes expressed in the PBT gonads at the sex-differentiated stage. Of these, 522 and 281 genes were upregulated in females and males, respectively. The qRT-PCR validated the expression profile and revealed the temporal expression patterns of the upregulated genes during gonadal sex differentiation in PBT. qRT-PCR validation targeted key factors in sex steroidogenesis and conservative factors and unreported factors in fish sex differentiation, which were enriched in multiple GO pathways. The expression profile of genes encoding key enzymes in estrogen and androgen biosynthesis, *aromatase* and *cyp11c1* were also validated through in situ hybridization (see



**Figure 5.** Expression patterns of sex steroid biosynthesis- and sex differentiation- related genes in sex-reversed (genotypic female/phenotypic male) Pacific bluefin tuna gonads. Genotypic female gonads were artificially masculinized via aromatase inhibitor administration<sup>1</sup>. Total RNA extracted from gonads at the morphologically sex-differentiated stage (70 days post-hatching) for each sex genotype in aromatase inhibitor-treated and untreated groups were subjected to quantitative real-time reverse transcription PCR. Bars indicate mean  $\pm$  standard error of the mean (SEM) ( $n = 3$  fish). Different letters indicate significant differences ( $P < 0.05$ ; two-way ANOVA followed by Tukey's multiple comparisons test).

Supplementary Fig. S3 online). Hence, this study provides numerous candidate genes potentially responsible for gonadal sex differentiation in PBT.

The enriched GO terms further characterized the upregulated genes for each sex. The genes upregulated in the females were highly enriched in cell/tissue development-related GO terms. These genes may be linked to ovarian cavity and lamellae formation because somatic cells proliferate in the female gonads to form them during ovarian differentiation<sup>28,29</sup>. We previously established the formation of these structures during ovarian differentiation in PBT<sup>6</sup>. By contrast, several upregulated genes in the males were enriched in structural development-related GO terms. These genes may be linked to the development of a structural framework, namely, the seminiferous epithelium, which is required for spermatogenesis in testes. The seminiferous epithelium is composed of Sertoli and germline cells<sup>30,31</sup>. The Sertoli cells create several junctions to provide developing germ cells with essential structural support<sup>30,31</sup>. Thus, our enrichment analysis disclosed numerous candidate genes responsible for developing ovarian and testicular structures in PBT.

Sex hormones play vital roles in gonadal sex differentiation in fish<sup>10</sup>. In this study, we found that the genes upregulated in the males were enriched in steroidogenesis-related GO terms. Furthermore, certain genes encoding key enzymes implicated in gonadal sex steroidogenesis, including *star*, *cyp11a2*, *cyp17a1*, *cyp17a2*, *cyp11c1*, and *hsd11b2*, were upregulated in the males. *Cyp11c1* and *hsd11b2* encode key enzymes in androgen biosynthesis,

and qRT-PCR confirmed that they were upregulated in differentiating male gonads. We previously confirmed that the males have a higher plasma androgen (11-ketotestosterone) concentration than the females at the gonadal sex-differentiated stage<sup>1</sup>. Our results indicate that androgen biosynthesis is upregulated in the male gonads during testicular differentiation in PBT. Exogenous androgens, such as 17- $\alpha$ -methyltestosterone, induce masculinization in various species of fish with the female genotype<sup>22</sup>. However, the roles of endogenous androgens in testicular differentiation in fish are poorly understood. Androgen biosynthesis is absent in differentiating male Nile tilapia gonads<sup>13,14</sup>. Therefore, endogenous androgens may not be essential for testicular differentiation in this species. In contrast, androgen biosynthesis is upregulated in differentiating male rainbow trout gonads<sup>15,16</sup>. Nevertheless, it only occurs after early testicular differentiation<sup>32</sup>. Hence, testicular differentiation may be initiated in the absence of endogenous androgens in this species. In male PBT gonads, *cyp11c1* and *hsd11b2* were upregulated only after the onset of morphological sex differentiation. Li et al.<sup>12</sup> suggested that endogenous androgens maintain testicular fate by suppressing ovarian differentiation in fish. In PBT, the upregulation of androgen biosynthesis may not directly induce testicular differentiation; rather, it might indirectly promote testicular differentiation by suppressing ovarian differentiation.

Unlike androgens, endogenous estrogens are essential for ovarian differentiation in fish<sup>10–12</sup>. We previously demonstrated *aromatase* upregulation and active estrogen biosynthesis in differentiating female PBT gonads<sup>1</sup>. We also showed that AI administration induced a sex reversal from genotypic female to phenotypic male<sup>1</sup>. The preceding results suggest that active aromatase-catalyzed estrogen biosynthesis is crucial for ovarian differentiation in PBT<sup>1</sup>, as reported in many fish species<sup>10–12</sup>. In the current study, our comparative transcriptome analysis revealed *foxl2* upregulation in the females. This gene is also critical in ovarian differentiation, which regulates the *aromatase* transcription either by directly binding its promoter or interacting with nuclear receptor subfamily 5 group A member 1 (Nr5a1)<sup>10–12</sup>. The qRT-PCR analysis disclosed that *foxl2* was upregulated mainly in differentiating female gonads. Notably, both *foxl2* and *aromatase* were more highly upregulated in the females than in the males at the morphologically sex-undifferentiated stage. Similar to prior reports on other fish species<sup>11,12</sup>, our results suggested that Foxl2 also plays a central role in ovarian differentiation in PBT by upregulating *aromatase*. Furthermore, downregulation of the genes controlling testicular differentiation was observed in *foxl2*-knockout genotypic female Nile tilapia<sup>33</sup>. Foxl2 may also contribute to ovarian differentiation in PBT by suppressing the genes regulating testicular differentiation. Future studies should be undertaken to confirm the role for Foxl2 in ovarian differentiation in tuna, particularly whether it transcriptionally regulates *aromatase* and the genes promoting testicular differentiation.

We isolated two genes from the *foxl2* family in PBT genome, *foxl2* (*foxl2a*) and *foxl3* (*foxl2b*; contig No. g15820). *Foxl3*, a paralog of *foxl2*, is essential for female germ cell fate decisions in medaka<sup>34</sup> and Nile tilapia<sup>35</sup>. In this study, comparative transcriptome analysis detected no *foxl3* expression in both sex PBT gonads at the sex-differentiated stage. Furthermore, no differences in germ cell development-related gene expression, including *vasa* (contig No. g25316), *dead end* (contig No. g24023), *nanos1* (contig No. g6249 and g10448), *nanos2* (contig No. g7592 and 7600), *syncp3* (*synaptonemal complex protein 3*; contig No. g29093), and *dmc1* (*DNA meiotic recombinase 1*; contig No. g15865) between sexes were observed (see Supplementary Tables S1 and 3 online). Our previous analysis revealed that PBT germ cells develop in a sexually dimorphic manner after sex differentiation: germ cell proliferation and differentiation occur earlier in the ovaries than in the testes<sup>6</sup>. Gene expressions in germ cells inducing sexual dimorphic development are expected to occur after gonadal sex differentiation in PBT. Notably, two contigs were predicted as genes encoding *nanos1* and *nanos2* from the updated PBT genome. In contrast, no contig was predicted as *nanos3*. Further studies are required to understand the existence of *nanos* family genes in the PBT genome.

Our comparative transcriptome analysis and qRT-PCR revealed equal expression levels of estrogen receptor encoding genes, *esr1*, *esr2a*, and *esr2b*, between sexes in the sex-differentiated stage PBT gonads, similar to the previous observations in several fish species, such as rainbow trout<sup>36</sup> and Nile tilapia<sup>14</sup>. The *esr2a* expression level was relatively high compared with those of *esr1* and *esr2b*, suggesting that Esr2a may primarily mediate estrogens in female PBT gonads during ovarian differentiation, although its role in male gonads during sex differentiation remains unclear.

20 $\beta$ -HSD plays a pivotal role in final oocyte maturation by producing MIH, i.e., 17 $\alpha$ , 20 $\beta$ -dihydroxy-4-pregnen-3-one<sup>26,27</sup>. In our comparative transcriptome analysis, no *hsd20b2* expression was detected in either sex. This is a reasonable result because no oocytes were observed in the sex-differentiated stage PBT ovaries subjected to RNA-seq.

The transcription factor Dmrt1 is implicated in vertebrate testicular differentiation<sup>2</sup>. Furthermore, Gsdf promotes testicular differentiation in fish<sup>10</sup>. Our comparative transcriptome analysis identified *dmrt1* and *gsdf* upregulation in the male PBT. Our qRT-PCR analysis revealed *dmrt1* and *gsdf* upregulation in both differentiating male and masculinized female (genotypic female/phenotypic male) gonads. The foregoing results and our understanding of the important roles of Dmrt1 and Gsdf in fish testicular differentiation<sup>10</sup> suggest that Dmrt1 and Gsdf regulate this developmental process in PBT.

AMH signaling comprises Amh and its receptor Amhr2, and it has a vital function in testicular differentiation in several fish species<sup>10</sup>. Furthermore, several studies have suggested that Sox9s (two types of *sox9* genes are found in fish, namely, *sox9a* and *sox9b* (or *sox9a2*), because of teleost-specific genome duplication<sup>37</sup>, whereas mammals possess only a single copy of *sox9*) may participate in gonadal sex differentiation in fish<sup>32,38</sup>. Nevertheless, our comparative transcriptome analysis showed that *amh*, *amhr2*, *sox9a*, and *sox9b* were expressed at the same levels in both sexes. AMH signaling and Sox9s are generally regarded as gonadal development factor and sex differentiation regulators in fish<sup>10</sup>. In PBT, these factors may be involved in ovarian and testicular development rather than gonadal sex differentiation.

In fish gonads, sexually dimorphic expression of the genes regulating sex differentiation generally occurs before morphological sex differentiation<sup>10</sup>. Our previous<sup>1</sup> and current studies confirmed *aromatase* and *foxl2*

upregulation in female PBT gonads at the morphologically sex-undifferentiated stage (41 dph). By that time, gonadal sex differentiation had already been initiated at the gene expression level. In contrast, the expression levels of *dmrt1* and *gsdf* did not differ between the sexes at this stage. We speculate that *Foxl2* and aromatase triggering ovarian differentiation, whereas *Dmrt1* and *Gsdf* are not required for triggering testicular differentiation. However, they are vital for promoting testicular differentiation. Incidentally, our qRT-PCR analysis identified upregulation of *cldn11a* in male PBT gonads at the morphologically sex-undifferentiated stage (41 dph). In mice, *claudin 11* expresses in Sertoli cells and is an obligatory protein for forming tight junctions in testis<sup>39</sup>. *Cldn11a* may be an upstream factor that initially induces testicular differentiation in fish. To the best of our knowledge, this is the first report on the relation of *cldn11a* to the early gonadal sex differentiation in fish. This gene will be studied in detail in the future.

We recently identified a male-specific homolog of *sulfotransferase family 1, cytosolic sulfotransferase 6* (*sult-1st6y*; contig No. g29404), as a candidate sex-determining gene in PBT<sup>25</sup>. Sulfotransferases deactivate endogenous estrogens by sulfating them<sup>40</sup>. Our comparative transcriptome analysis revealed *sult1st6y* upregulation in the males (see Supplementary Table S2 online). Furthermore, we previously demonstrated that inhibiting estrogen biosynthesis via AI treatment induces testicular differentiation in female gonads<sup>1</sup>. We speculated that *Sult1st6y* would trigger testicular differentiation through estrogen deactivation in male gonads at the onset of gonadal sex differentiation<sup>25</sup>, and estrogen deficiency induces upregulation of the genes promoting testicular differentiation, particularly *dmrt1* and *gsdf*. We are currently investigating *sult1st6y* expression patterns and the mechanism by which *Sult1st6y* deactivates estrogens.

In conclusion, we have developed a global gene expression profile for gonadal sex differentiation in PBT using NGS technology. Based on our findings, we propose that ovarian differentiation is mainly induced by aromatase and *Foxl2*, whereas *Dmrt1* and *Gsdf* play central roles in testicular differentiation in PBT. We previously demonstrated the morphological characteristics of gonadal sex differentiation and the vital role of aromatase-mediated estrogen synthesis in this developmental process in PBT<sup>1,6</sup>. Overall, the discoveries made herein lay theoretical and empirical foundations for understanding the sex differentiation mechanisms in PBT and other tuna species.

## Methods

**Ethics statement.** All experiments were conducted in accordance with the Guidelines for the Care and Use of Live Fish of the Fisheries Technology Institute (FTI) of the Japan Fisheries Research and Education Agency (FRA) and the ARRIVE guidelines. All experiments were approved by the Institutional Animal Care and Use Committee of FTI.

**Draft genome improvement and protein coding sequence prediction.** The draft genome sequence for male PBT consisted of 948 scaffolds and was previously published in 2021<sup>25</sup>. This sequence was polished with PolishCLR<sup>41</sup> (<https://github.com/isugifNF/polishCLR>) using the Pacific Biosciences (PacBio) and Illumina data from the previous version. Genome completeness was evaluated and compared against that of the previous version using BUSCO<sup>42</sup> v. 5.3.0 (<https://gitlab.com/ezlab/busco/-/releases#5.3.0>) and the Actinopterygii ortholog set. The protein-coding sequences were predicted with reference to those from the PBT genome published in 2013<sup>23</sup>, the genome of Southern bluefin tuna (*Thunnus maccoyii*) in the GenBank<sup>43</sup> (<https://www.ncbi.nlm.nih.gov/genbank/>), and five model fish genomes in the Ensembl database<sup>44</sup> (release 104) (<https://www.ensembl.org/index.html?redirect=n0>) including zebrafish, stickleback (*Gasterosteus aculeatus*), medaka, tiger puffer (*Takifugu rubripes*), and tetraodon (*Tetraodon nigroviridis*). The foregoing protein sequences were mapped to the current PBT genome with Exonerate<sup>45</sup> v. 2.4.0 (<https://anaconda.org/bioconda/exonerate>). All prediction results were merged excluding the overlapping loci with lower alignment scores. The predicted protein sequences were compared with those of two model fishes, zebrafish and medaka (Ensembl database, Release 104), using the BLASTP program (E-value < 1.0E<sup>-4</sup>).

**Sample collection and RNA extraction and sequencing.** PBT gonads at the sex-differentiated stage (Fig. 1a) were collected at 100 dph from fish reared from fertilized eggs. The oogonia and type-A spermatogonia were solitarily distributed in the differentiated ovary and testis. The total length and body weight of all PBTs sampled were 25.6 ± 11.5 cm and 258.4 ± 94.5 g ( $n = 10$  fish, mean ± standard error of the mean (SEM)), respectively. Dissected gonads were stored in RNAlater (Ambion, Austin, TX, USA) at -30 °C until RNA extraction. The sex phenotypes and genotypes of each fish were determined via histological observation and PCR-based genotypic sex identification<sup>6,25</sup>. All fish gonads differentiated into ovaries or testes according to their sex genotypes.

Total RNA was extracted from the gonads, and RNA extracts were treated with DNase using the RNeasy Mini kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The concentration and quality of purified total RNA were determined using NanoDrop One (Thermo Fisher Scientific, Waltham, MA) and the Agilent 2200 TapeStation System with RNA screen tape (Agilent Technologies, Santa Clara, CA). High-quality RNA extracts with RNA integrity number (RIN) > 7 were selected from three fish per sex (Table 1). The cDNA libraries were constructed using 1 µg total RNA and Illumina Stranded mRNA Prep (Illumina, San Diego, CA) according to the manufacturer's instructions. The libraries were sequenced into 75-bp paired end reads using NextSeq 500 systems (Illumina). The raw data were demultiplexed, trimmed of adapters, and converted into the FASTQ format using bcl2fastq v. 2.20 (Illumina).

**Comparative transcriptome analysis.** RNA-Seq data analysis was performed using the CLC Genomics Workbench v. 12.0.3 (QIAGEN GmbH, Hilden, Germany). Raw reads were trimmed with default parameters (quality limit = 0.05 and ambiguous limit = 2), and the trimmed reads were mapped onto the reference sequences

with default parameters (mismatch cost = 2, insertion cost = 3, deletion cost = 3, length fraction = 0.8, and similarity fraction = 0.8). The read count files containing the values of the reads mapped against the reference sequences were then loaded into the integrated Differential Expression and Pathway analysis (iDEP) web platform<sup>46</sup> (<http://bioinformatics.sdstate.edu/idep/>). The data were then normalized using DESeq2<sup>47</sup> (<https://bioconductor.org/packages/release/bioc/html/DESeq2.html>). The DESeq2 program in iDEP was also used to identify differentially expressed genes between sexes. The false discovery rate and minimum fold change cutoff/threshold values were set to 0.05 and 2, respectively. The differentially expressed genes were then subjected to enrichment analysis in the GO biological process database using the Ensembl zebrafish gene IDs (Release 104) in iDEP.

**qRT-PCR.** Previously sampled PBT gonads<sup>1,6</sup> were stored in RNAlater (Ambion) at  $-30\text{ }^{\circ}\text{C}$  until RNA extraction. Total RNA extraction and qRT-PCR were performed as previously described<sup>1</sup>. The qRT-PCR reaction volumes contained  $0.4\text{ }\mu\text{M}$  of each forward and reverse gene-specific primers, except for *anxa1b*, which contained  $0.1\text{ }\mu\text{M}$  of each forward and reverse primer. The gene-specific primers and PCR amplification conditions are shown in Supplementary Table S6. The primers were designed using CLC Main Workbench v. 8.0.1 (QIAGEN GmbH). The quantification standard was a plasmid containing the partial cDNA sequence of a target gene. Standard seven-point sets ranged from  $1 \times 10^8$  to  $1 \times 10^2$  copies and were prepared using tenfold serial dilutions. Technical duplicates were performed for all experimental samples and standards. The intra-assay coefficient of variation (CV) was determined using repeated standard sample measurements ( $n=8$ ) (see Supplementary Table S6 online). The expression levels are presented as means  $\pm$  SEM.

**Statistical analysis.** For the comparative transcriptome analysis, statistical analyses were conducted in iDEP. For the temporal gene expression comparison using qRT-PCR, the statistical analyses included Welch's *t*-test (between sexes at each sampling period) and one-way analysis of variance (ANOVA), followed by Tukey's multiple comparisons test (for the expression trends in each sex). Statistical analyses of the gene expression levels in the sex-reversed gonads included two-way ANOVA followed by Tukey's multiple comparisons test. Statistical analyses of the qRT-PCR data were performed in GraphPad Prism v. 7.0 (GraphPad Software, San Diego, CA, USA). Statistically significant differences were determined at  $P < 0.05$ .

### Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The polished genome sequences were deposited in the DNA Data Bank of Japan (DDBJ) under Accession Numbers BOUD02000001–BOUD02000580. The RNA sequencing results were deposited in the DDBJ Sequence Read Archive (DRA) under Accession Numbers DRR458350–DRR458355.

Received: 4 May 2023; Accepted: 18 August 2023

Published online: 24 August 2023

### References

- Hayashida, T. *et al.* First data on sexual dimorphic growth of cultured Pacific bluefin tuna, *Thunnus orientalis* (Temminck et Schlegel), and its sex manipulation by treatment with an aromatase inhibitor. *Aquac. Res.* **2023**, 1590646. <https://doi.org/10.1155/2023/1590646> (2023).
- Shimose, T. & Farley, J. H. Age, growth and reproductive biology of bluefin tunas. In *Biology and Ecology of Bluefin Tuna* (eds Kitagawa, T. & Kimura, S.) 47–77 (CRC Press, 2015).
- Gen, K. Physiology of bluefin tuna reproduction. In *Biology and Ecology of Bluefin Tuna* (eds Kitagawa, T. & Kimura, S.) 325–354 (CRC Press, 2015).
- Zohar, Y., Mylonas, C. C., Rosenfeld, H., de la Gándara, F. & Corriero, A. Reproduction, broodstock management, and spawning in captive Atlantic bluefin tuna. In *Advances in Tuna Aquaculture From Hatchery to Market* (eds Benetti, D. D. *et al.*) 159–188 (Academic Press, 2016).
- Higuchi, K. *et al.* Early development of primordial germ cells in Pacific bluefin tuna *Thunnus orientalis*. *Theriogenology* **131**, 106–112 (2019).
- Hayashida, T. *et al.* Gonadal sex differentiation and early ovarian/testicular development in cultured Pacific bluefin tuna, *Thunnus orientalis* (Temminck et Schlegel). *Theriogenology* **173**, 56–63 (2021).
- Higuchi, K. *et al.* Seasonal changes and endocrine regulation of gonadal development in hatchery-produced Pacific bluefin tuna *Thunnus orientalis* broodstock in sea cages. *Aquaculture* **545**, 737199. <https://doi.org/10.1016/j.aquaculture.2021.737199> (2021).
- Hayashida, T. *et al.* Expression of the *dead end* gene during embryogenesis and its role in primordial germ cell development in Pacific bluefin tuna, *Thunnus orientalis* (Temminck et Schlegel). *Aquac. Res.* **53**, 3872–3876 (2022).
- Devlin, R. H. & Nagahama, Y. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture* **208**, 191–364 (2002).
- Nagahama, Y., Chakraborty, T., Paul-Prasanth, B., Ohta, K. & Nakamura, M. Sex determination, gonadal sex differentiation, and plasticity in vertebrate species. *Physiol. Rev.* **101**, 1237–1308 (2021).
- Guiguen, Y., Fostier, A., Piferrer, F. & Chang, C.-F. Ovarian aromatase and estrogens: A pivotal role for gonadal sex differentiation and sex change in fish. *Gen. Comp. Endocrinol.* **165**, 352–366 (2010).
- Li, M., Sun, L. & Wang, D. Roles of estrogens in fish sexual plasticity and sex differentiation. *Gen. Comp. Endocrinol.* **277**, 9–16 (2019).
- Nakamura, M., Bhandari, R. K. & Higa, M. The role estrogens play in sex differentiation and sex changes of fish. *Fish Physiol. Biochem.* **28**, 113–117 (2003).
- Ijiri, S. *et al.* Sexual dimorphic expression of genes in gonads during early differentiation of a teleost fish, the Nile Tilapia *Oreochromis niloticus*. *Biol. Reprod.* **78**, 333–341 (2008).
- Liu, S. J. *et al.* Expression of cytochrome P45011 $\beta$  (11 $\beta$ -hydroxylase) gene during gonadal sex differentiation and spermatogenesis in rainbow trout, *Oncorhynchus mykiss*. *J. Steroid Biochem. Mol. Biol.* **75**, 291–298 (2000).
- Govoroun, M. *et al.* Steroid enzyme gene expressions during natural and androgen-induced gonadal differentiation in the rainbow trout. *Oncorhynchus Mykiss*. *J. Exp. Zool.* **290**, 558–566 (2001).
- Webster, K. A. Dmrt1 is necessary for male sexual development in zebrafish. *Dev. Biol.* **422**, 33–46 (2017).

18. Wei, L. *et al.* Dmrt1 directly regulates the transcription of the testis-biased *Sox9b* gene in Nile tilapia (*Oreochromis niloticus*). *Gene* **687**, 109–115 (2019).
19. Wang, D.-S. *et al.* Doublesex- and Mab-3-related transcription factor-1 repression of aromatase transcription, a possible mechanism favoring the male pathway in tilapia. *Endocrinology* **151**, 1331–1340 (2010).
20. Kaneko, H. *et al.* Gonadal soma-derived factor (*gsdf*), a TGF-beta superfamily gene, induces testis differentiation in the teleost fish *Oreochromis niloticus*. *Mol. Cell. Endocrinol.* **415**, 87–99 (2015).
21. Zhang, X. *et al.* Autosomal *gsdf* acts as a male sex initiator in the fish medaka. *Sci. Rep.* **6**, 19738. <https://doi.org/10.1038/srep19738> (2016).
22. Budd, A. M., Banh, Q. Q., Domingos, J. A. & Jerry, D. R. Sex control in fish: Approaches, challenges and opportunities for aquaculture. *J. Mar. Sci. Eng.* **3**, 329–355 (2015).
23. Nakamura, Y. *et al.* Evolutionary changes of multiple visual pigment genes in the complete genome of Pacific bluefin tuna. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11061–11066 (2013).
24. Suda, A. *et al.* Improvement of the Pacific bluefin tuna (*Thunnus orientalis*) reference genome and development of male-specific DNA markers. *Sci. Rep.* **9**, 14450. <https://doi.org/10.1038/s41598-019-50978-4> (2019).
25. Nakamura, Y. *et al.* Prediction of the sex-associated genomic region in tunas (*Thunnus* fishes). *Int. J. Genom.* **2021**, 7226353. <https://doi.org/10.1155/2021/7226353> (2021).
26. Young, G., Kusakabe, M., Nakamura, I., Lokman, P. M. & Goetz, F. W. Gonadal steroidogenesis in teleost fish. In *Hormones and Their Receptors in Fish Reproduction* (eds Melamed, P. & Sherwood, N.) 155–223 (World Scientific, 2005).
27. Rajakumar, A. & Senthilkumaran, B. Steroidogenesis and its regulation in teleost—a review. *Fish Physiol. Biochem.* **46**, 803–818 (2020).
28. Nakamura, M., Kobayashi, T., Chang, X.-T. & Nagahama, Y. Gonadal sex differentiation in teleost fish. *J. Exp. Zool.* **281**, 362–372 (1998).
29. Strüssmann, C. A. & Nakamura, M. Morphology, endocrinology, and environmental modulation of gonadal sex differentiation in teleost fishes. *Fish Physiol. Biochem.* **26**, 13–29 (2002).
30. Vogl, A. W., Pfeiffer, D. C., Mulholland, D., Kimel, G. & Guttman, J. Unique and multifunctional adhesion junctions in the testis: Ectoplasmic specializations. *Arch. Histol. Cytol.* **63**, 1–15 (2000).
31. Batlouni, S. R., Nóbrega, R. H. & França, L. R. Cell junctions in fish seminiferous epithelium. *Fish Physiol. Biochem.* **35**, 207–217 (2009).
32. Vizziano, D., Randuineau, G., Baron, D., Cauty, C. & Guiguen, Y. Characterization of early molecular sex differentiation in rainbow trout. *Oncorhynchus Mykiss*. *Dev. Dyn.* **236**, 2198–2206 (2007).
33. Zhang, X. *et al.* Mutation of *foxl2* or *cyp19a1a* results in female to male sex reversal in XX Nile tilapia. *Endocrinology* **158**, 2634–2647 (2017).
34. Nishimura, T. *et al.* *foxl3* is a germ cell–intrinsic factor involved in sperm-egg fate decision in medaka. *Science* **349**, 328–331 (2015).
35. Dai, S. *et al.* Germline sexual fate is determined by the antagonistic action of *dmrt1* and *foxl3/foxl2* in tilapia. *Development* **148**, dev199380. <https://doi.org/10.1242/dev.199380> (2015).
36. Guiguen, Y. *et al.* Involvement of estrogens in the process of sex differentiation in two fish species: The rainbow trout (*Oncorhynchus mykiss*) and a tilapia (*Oreochromis niloticus*). *Mol. Reprod. Dev.* **54**, 154–162 (1999).
37. Klüver, N., Kondo, M., Herpin, A., Mitani, H. & Schartl, M. Divergent expression patterns of Sox9 duplicates in teleosts indicate a lineage specific subfunctionalization. *Dev. Genes Evol.* **215**, 297–305 (2005).
38. Wei, L., Yang, C., Tao, W. & Wang, D. Genome-wide identification and transcriptome-based expression profiling of the Sox gene family in the Nile tilapia (*Oreochromis niloticus*). *Int. J. Mol. Sci.* **17**, 270. <https://doi.org/10.3390/ijms17030270> (2016).
39. Mazaud-Guittot, S. *et al.* Claudin 11 Deficiency in mice results in loss of the Sertoli cell epithelial phenotype in the testis. *Biol. Reprod.* **82**, 202–213 (2010).
40. Negishi, M. *et al.* Structure and function of sulfotransferases. *Arch. Biochem. Biophys.* **390**, 149–157 (2001).
41. Rhie, A. *et al.* Towards complete and error-free genome assemblies of all vertebrate species. *Nature* **592**, 737–746 (2021).
42. Manni, M., Berkeley, M. R., Seppy, M. & Zdobnov, E. M. BUSCO: Assessing genomic data quality and beyond. *Curr. Protoc. I.* **323**. <https://doi.org/10.1002/cpz1.323> (2021).
43. Sayers, E. W. *et al.* GenBank. *Nucleic Acids Res.* **48**, 84–86 (2020).
44. Howe, K. L. *et al.* Ensembl 2021. *Nucleic Acids Res.* **49**, D884–D891 (2021).
45. Slater, G. S. & Birney, E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinform.* **6**, 31. <https://doi.org/10.1186/1471-2105-6-31> (2005).
46. Ge, S. X., Son, E. W. & Yao, R. iDEP: An integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinform.* **19**, 534. <https://doi.org/10.1186/s12859-018-2486-6> (2018).
47. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. <https://doi.org/10.1186/s13059-014-0550-8> (2014).

## Acknowledgements

This study was financially supported by the Fisheries Technology Institute (FTI), the Japan Fisheries Research and Education Agency (FRA), and a Grant from the Japan Society for the Promotion of Science (JSPS) KAKENHI (No. JP21H02275).

## Author contributions

T.H., K.H., Y.K., and K.G. designed the study. T.H. and S.S. performed the RNA sequencing and transcriptomic analysis. Y.N. updated the draft genome and predicted the genes from the refined genome. T.H. performed the qRT-PCR analysis. T.H., S.S., Y.N., and Y.K. wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-40914-y>.

**Correspondence** and requests for materials should be addressed to T.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023